

U.S. Decennial Life Tables for 2009–2011, Methodology of the United States Life Tables

by Elizabeth Arias, Ph.D., Sally G. Curtin, M.A., and Betzaida Tejada-Vera, M.S., Division of Vital Statistics

Abstract

Objectives—This report describes the methodology used in the preparation of the 2009–2011 decennial life tables for the United States by race, Hispanic origin, and sex based on the age-specific death rates for the period 2009–2011, appearing in the report, “U.S. Decennial Life Tables for 2009–2011, United States Life Tables” (1).

Methods—Data used to prepare these life tables include population data by age on the census date April 1, 2010; deaths occurring in the 3-year period 2009–2011 classified by age at death; births for each of the years 2007–2011; and Medicare data for ages 66–99 for the 3 years 2009–2011. The methods used differ from those applied to the 1999–2001 decennial life tables in the estimation of mortality for ages 66 and over. For the total, white, black, non-Hispanic white, and non-Hispanic black populations, the method developed for the U.S. annual life tables beginning with data year 2008 was used. It consists of the application of the Kannisto logistic model to smooth death rates in the age range 85–99 and predict death rates for ages 100–120 (2,3). For the Hispanic population, which is added to the decennial series for the first time with the 2009–2011 set, the method developed for the U.S. annual life tables beginning with data year 2006 was used. This method consists of using the Brass relational logit model to estimate mortality for ages 80–120 (4).

Keywords: life expectancy • survival • death rates • race • Hispanic origin • National Vital Statistics System

Introduction

This report describes the methodology used in preparing the 2009–2011 decennial life tables for the United States. The following life tables were produced for the United States total population and for the male, female, white, white male, white female, black, black male, black female, Hispanic, Hispanic male, Hispanic female, non-Hispanic white, non-Hispanic white male,

non-Hispanic white female, non-Hispanic black, non-Hispanic black male, and non-Hispanic black female populations. These life tables are presented in the report, “U.S. Decennial Life Tables for 2009–2011, United States Life Tables” (1).

There are two types of life tables—the cohort (or generation) life table and the current (or period) life table. The cohort life table presents the mortality experience of a particular birth cohort from the moment of birth through consecutive ages in successive calendar years. For example, a cohort life table could be based on the mortality experience of all persons born in the year 1900 and followed until the cohort is extinct. Based on age-specific death rates observed through consecutive calendar years, the cohort life table reflects the mortality experience of an actual cohort from birth until no lives remain in the group.

The current life table may, by contrast, be characterized as cross-sectional. Unlike the cohort life table, the current life table considers a hypothetical (or synthetic) cohort that is subject throughout its existence to the age-specific mortality rates observed for an actual population during a particular period of relatively short duration (often 1–3 years). The current life table may thus be characterized as rendering a snapshot of current mortality experience and shows the long-range implications of a set of age-specific death rates that prevailed in a given year or period. The life tables presented in this report are current life tables based on age-specific death rates for the period 2009–2011.

The decennial life tables differ in two respects from the life tables prepared and published annually in the *National Vital Statistics Reports*. The annual tables are based on deaths in a single year and, except for the census year, on postcensal population estimates rather than on the enumerated population from a decennial census. The decennial tables are based on deaths in a period of 3 years, including 2009, 2010, and 2011, and on population counts from the 2010 decennial census.

The methods used to estimate the 2009–2011 decennial life tables differ from those applied to the 1999–2001 decennial life tables in the estimation of mortality for older ages. The methodology was refined in two important ways. First, a logistic



model rather than a nonlinear least squares model was used to smooth and extrapolate the vital statistics and Medicare-blended death rates at the oldest ages. Second, the age at which smoothing is begun was raised from 66 to 85 or so, depending on the population (2). This modification applies to the life tables for the total population and for the white, black, non-Hispanic white, and non-Hispanic black populations. The 1999–2001 set did not include life tables by Hispanic origin. Life tables by Hispanic origin, including Hispanic, non-Hispanic white, and non-Hispanic black, were first published with 2006 annual mortality data. For the Hispanic population, the method used was developed for the U.S. annual life tables beginning with data year 2006. This method consists of using the Brass relational logit model to estimate mortality for ages 80–120 (4).

Methods

Data used for calculating life table values

The data used in preparing each of the 2009–2011 decennial life tables consisted of: a) deaths occurring in a 3-year (2009–2011) period classified by age at death and collected from death certificates filed in state vital statistics offices and reported to the National Center for Health Statistics (NCHS) as part of the National Vital Statistics System (NVSS); b) population data by age on the census date April 1, 2010; c) Medicare data for deaths and population for ages 66–99 during a 3-year (2009–2011) period; and d) births for each of the years from 2007–2011.

Vital statistics data

Death counts used for computing the life tables presented in this report are final numbers of deaths for 2009–2011 collected from death certificates filed in state vital statistics offices and reported to NCHS as part of NVSS. Race and Hispanic origin are reported separately on the death certificate.

The U.S. Standard Certificate of Death was revised in 2003, and its race and Hispanic-origin items reflect the mandate of the 1997 Office of Management and Budget (OMB) standards (5). This revision allowed individuals to report more than one race and increased the race choices from four to five by separating the Asian and Pacific Islander groups. Because of the staggered implementation of the 2003 death certificate revision, national implementation of the 1997 standards was not complete until 2018. During 2009–2011, approximately 32 states and the District of Columbia had adopted the 1997 standards. To attain uniformity and comparability during the transition period, multiple-race responses were bridged back to the 1977 single-race standard, and Asian and Pacific Islander groups were combined according to the 1977 OMB standards. The bridging procedure was the same as that used to bridge multiple-race population estimates, as discussed in the next section (6).

Census population data

The census populations used in the denominators of the death rates are based on special estimation procedures and are not true counts of the 2010 decennial census. These estimates were produced under a collaborative agreement between NCHS and the U.S. Census Bureau. Reflecting the guidelines issued in 1997 by OMB, the 2010 census included an option for individuals to report more than one race as appropriate for themselves and household members (5). The 1997 OMB guidelines also provided for the reporting of Asian persons separately from the Native Hawaiian or Other Pacific Islander group. Under the prior OMB standards issued in 1977, data for Asian or Pacific Islander persons were collected as a single group (7). To produce death rates for the 2009–2011 period, it was necessary to bridge the reported population data for multiple-race persons back to single-race categories. In addition, the 2010 census counts were modified to be consistent with the 1977 OMB race categories—that is, to report the data for Asian persons and the Native Hawaiian or Other Pacific Islander group as a combined category, Asian or Pacific Islander. The procedures used to produce the bridged populations are described in separate publications (6).

With respect to census population data, actuarial theory suggests that the populations to be used in the calculations should be those of the central date of the 3-year period, that is, July 1, 2010. However, the enumerated populations as of April 1, 2010, were used as if they were July 1 populations. This was done because the percent differences between the two sets of population figures are very small at the national level. It was not considered necessary to produce new estimates to reflect the lapse of time between April 1 and July 1, 2010.

Medicare data

Medicare data have been employed in the estimation of U.S. decennial life tables beginning with the 1969–1971 set (8). Medicare data are considered more accurate than vital statistics and census data at the oldest ages because Medicare enrollees must have proof of age in order to enroll (9). However, the reliability of Medicare data beyond age 100 declines because of the small percentage of persons who enrolled at the start of the Medicare program in 1965 and for whom it was not possible to verify exact age (9). Further, the Medicare race and ethnicity classification system makes it impossible to correctly identify the Hispanic, American Indian or Alaska Native, or Asian or Pacific Islander populations (4,10). It is, however, possible to use Medicare data to estimate old-age mortality for both the white and black race groups, regardless of Hispanic origin, as has been done traditionally, and to estimate old-age mortality for the non-Hispanic segments of these populations (4). As a result, data from the Medicare program are used to supplement vital statistics and census data for ages 66–99 for the total population and for the white, black, non-Hispanic white, and non-Hispanic black populations (4).

To estimate death rates for the Medicare white, black, non-Hispanic white, and non-Hispanic black populations in 2009–2011, age-specific numbers of deaths and population counts by sex and race for the population aged 66–99 from the

2009–2011 Medicare files were used. The data files are created by the Centers for Medicare & Medicaid Services for the Social Security Administration, which under a special agreement shares the data with NCHS. The 2009–2011 files contain midyear Medicare population counts (as of June 30) and calendar year Medicare death counts (January 1–December 31) for each year of 2009–2011.

Preliminary adjustment of the data

Adjustment for unknown age

An adjustment is made to account for the small proportion of deaths for each year during 2009–2011 for which age is not reported on the death certificate. The assumption was made that these deaths were distributed among the various age groups in the same proportions as the deaths for which age was reported. To this end, adjustment factors were computed for each population category for which a life table was to be constructed. This factor was obtained by dividing the total number of deaths reported for the given category for the 3-year period 2009–2011 by the total number of deaths with age stated. That is,

$$F = \frac{D}{D^a},$$

where D is the total number of deaths and D^a is the total number of deaths for which age is stated. Death records with missing age information in each given Hispanic-origin, race, and sex category were then proportionally distributed among age categories as

$$D_x^{new} = F \cdot D_x^{original}.$$

Adjustment for misclassification of race and Hispanic origin on death certificates

The latest research to evaluate race and Hispanic-origin reporting on U.S. death certificates found that misclassification of race and Hispanic origin on death certificates in the United States accounts for a net underestimate of 3% for total Hispanic deaths, a net underestimate of less than one-half percent for total non-Hispanic black deaths, and no under- or overestimate for total non-Hispanic white deaths or for the population racially classified as white or black, regardless of Hispanic origin (11). These results are based on a comparison of self-reported race and Hispanic origin on Current Population Surveys (CPS) with race and Hispanic origin reported on the death certificates of a sample of decedents in the National Longitudinal Mortality Study (NLMS) who died during the period 1999–2011 (11).

NLMS-linked records are used to estimate sex-age-specific ratios of CPS race and Hispanic-origin counts to death certificate counts (11–13). The CPS/death certificate ratio, or classification ratio, is specifically the ratio of the weighted count of self-reported race and ethnicity on the CPS to the weighted count of the same racial or ethnic category on the death certificates of the sample of NLMS decedents described above. It can be interpreted as the net difference in assignment of a specific race and Hispanic-origin category between the two classification systems and can be used as a correction factor for race and Hispanic-origin

misclassification (11–13). The assumption is made that the race and ethnicity reported by a CPS respondent is more reliable than proxy reporting of race and ethnicity by a funeral director who has little personal knowledge of the decedent. Further, public policy embodied in the 1997 OMB standard mandates that self-identification should be the standard used for the collection and recording of race and ethnicity information (5).

The NLMS-based classification ratios discussed above are used to adjust the age-specific number of deaths for ages 1–95 and over for the total Hispanic, non-Hispanic white, and non-Hispanic black populations, and by sex for each group, as follows:

$${}_nD_x = {}_nD_x^F \cdot {}_nCR_x,$$

where ${}_nD_x^F$ is the age-specific number of deaths adjusted for unknown age as described above, ${}_nCR_x$ are the sex- and age-specific classification ratios used to correct for the misclassification of race and Hispanic origin on death certificates, and ${}_nD_x$ are the final age-specific counts of death adjusted for age and race and Hispanic-origin misclassification. Table A shows values of the sex- and age-specific classification ratios, ${}_nCR_x$, by Hispanic origin and race for the non-Hispanic population (both black and white).

Because NLMS classification ratios for infant deaths are unreliable due to small sample sizes, corrections for racial and ethnic misclassification of infant deaths are addressed by using infant death counts and live birth counts from the 2007–2011 linked birth/infant death data files rather than the traditional birth and death data files (14–18). In the linked file, each infant death record is linked to its corresponding birth record so that the race and ethnicity of the mother reported on the birth record can be ascribed to the infant death record. As a result, race- and ethnicity-specific infant mortality rates estimated with the linked file do not suffer from the problem of racial and ethnic discrepancies between the numerator and denominator of the rate.

Note that although no conclusive evidence supports return migration as a factor in the lower mortality of the Hispanic population, the possibility remains that Hispanic deaths are missed in NVSS due to return migration, and therefore, the resulting death rates may be biased regardless of correction for ethnic misclassification (4,19).

In preparing the 2009–2011 decennial life tables, no specific allowance was made for possible incompleteness in the enumeration of the population or in the registration of births or deaths. In calculating previous decennial life tables, the use of birth statistics rather than population data in calculating the denominators of the mortality rates at ages under 2 years was justified largely on the basis that the census populations under 2 years of age were believed to be underenumerated. In addition, using the methodology based on birth data had other advantages because it might be expected to produce a more accurate estimate of the average population during the 3-year period than is provided by the population enumerated on the census date. Accordingly, its use was continued in the 2009–2011 life tables.

Table A. Classification ratios, by Hispanic origin, race for the non-Hispanic white and black populations, age, and sex

Age (years)	Hispanic			Non-Hispanic white			Non-Hispanic black		
	Total	Male	Female	Total	Male	Female	Total	Male	Female
All ages	1.0329	1.0362	1.0294	0.9995	0.9993	0.9997	1.0047	1.0041	1.0053
1–14	0.9905	0.9659	*1.0299	0.9918	1.0755	0.8770	1.0266	0.9379	*1.1751
15–24	0.9668	0.9325	1.0604	0.9976	1.0019	0.9869	1.0248	1.0215	1.0343
25–34	1.0354	1.0401	1.0232	1.0021	1.0034	0.9994	0.9855	0.9770	1.0008
35–44	1.0434	1.0645	1.0066	0.9980	0.9997	0.9951	1.0062	1.0073	1.0048
45–54	1.0584	1.0372	1.0953	0.9969	0.9965	0.9976	1.0002	1.0019	0.9982
55–64	1.0571	1.0517	1.0659	0.9994	0.9992	0.9997	1.0003	0.9965	1.0046
65–74	1.0295	1.0485	1.0072	0.9967	0.9967	0.9966	1.0062	1.0055	1.0070
75–84	1.0192	1.0188	1.0196	1.0004	1.0003	1.0004	1.0057	1.0057	1.0058
85–94	1.0208	1.0313	1.0137	1.0008	1.0007	1.0009	1.0110	1.0155	1.0086
95 and over	1.0732	1.0509	1.0842	1.0005	0.9995	1.0008	0.9980	0.9872	0.9954

* Ratio does not meet National Center for Health Statistics standards of reliability because either the unweighted number of Current Population Survey deaths or the unweighted number of death certificate deaths, or both, are based on fewer than 20 deaths.

SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

Populations and deaths were available by single years of age from 0 through 100 and over. In each case, the age referred to is in completed years—that is, the exact age on the individual’s last birthday. In addition, deaths occurring at ages under 1 year were available for four subdivisions of the first year of life: under 1 day, 1–7 days, 7–28 days, and 28–365 days. Life table values were calculated for these age groups within the first year of life and for single years of age throughout the remainder of the life span.

Calculation of the probability of dying, ${}_nq_x$

The life table function ${}_nq_x$ is the fraction or proportion of a group of persons at exact age x who are expected to die before attaining age $x + n$. The ${}_nq_x$ is also called the probability of dying. Other functions in the complete life table are derived from ${}_nq_x$, which depends on the number of deaths ${}_nD_x$ and the midyear population ${}_nP_x$ for each age interval x to $x + n$ observed during the calendar year of interest.

In constructing the decennial life tables, ${}_nD_x$ for years 2009–2011 and ${}_nP_x$ for decennial year 2010 were smoothed using Beer’s graduation technique (20). Beer’s multipliers were applied to observed deaths in 5-year age intervals to get smoothed single-year-of-age data. Similarly, Beer’s multipliers were applied to population estimates in 5-year age intervals to get smoothed estimates for single-year age intervals.

For ages under 2 years, births, B , instead of ${}_nP_x$ were used to calculate ${}_nq_x$. Life table deaths, ${}_nd_x$, and surviving populations at the beginning of each age, l_x , were derived from ${}_nq_x$.

Probabilities of dying at ages under 2 years

At ages under 2 years, the first life table quantities to be calculated were the values of ${}_nd_x$, the number of deaths occurring between exact age x and $x + n$ in the life table cohort commencing with l_0 live births. This was calculated by the formula

$${}_nd_x = l_0 \cdot \frac{{}_nD_x}{{}_nE_x},$$

where ${}_nd_x$ is the estimated number of deaths per the life table cohort of 100,000 live births, l_0 , between exact ages x and $x + n$, ${}_nD_x$ is the actual number of deaths (adjusted for age, race, and Hispanic-origin misreporting) occurring in 2009–2011, and ${}_nE_x$ is the appropriate denominator as indicated in Table B. The denominators, ${}_nE_x$, are based on the assumption that births in 2007, 2008, 2009, 2010, and 2011 are uniformly distributed over each year. Age intervals for under 2 years were classified as 0 days (under 1 day), 1–7 days, 7–28 days, 28–365 days, and 0–1 year. The number of deaths at age 0 is the sum of deaths occurring at ages 0 to 364 days. The number of deaths at age 1 year includes all deaths occurring from 365 days to the day just before the second birthday.

Table B. Denominators ${}_nE_x$ used in calculating ${}_nd_x$ for ages under 2 years: U.S. Decennial Life Tables, 2009–2011

Age interval x to $x + n$	${}_nE_x$
Under 1 day	1/730 ($B_{2008} + 730B_{2009} + 730B_{2010} + 729B_{2011}$)
1–7 days	1/730 ($8B_{2008} + 730B_{2009} + 730B_{2010} + 722B_{2011}$)
7–28 days	1/730 ($35B_{2008} + 730B_{2009} + 730B_{2010} + 695B_{2011}$)
28–365 days	1/730 ($393B_{2008} + 730B_{2009} + 730B_{2010} + 337B_{2011}$)
1–2 years	1/2 ($B_{2007} + 2B_{2008} + 2B_{2009} + B_{2010}$)

SOURCE: National Center for Health Statistics, National Vital Statistics System, Natality.

The unrounded values of ${}_nd_x$ were then used to calculate the number of survivors, l_x , up to age 2 years by successive applications of the formula

$$l_{x+n} = l_x - {}_nd_x.$$

The probability of dying within each age category for the ages under 2 years was estimated as ${}_nd_x$ divided by l_x

$${}_nq_x = \frac{{}_nd_x}{l_x}.$$

Death rates at ages 2–94 from vital statistics data

Interpolation of P_x and D_x at ages 2–94

Anomalies, both random and those associated with age reporting, can be problematic when using vital statistics and census data by single years of age (20). Graduation techniques are often used to eliminate these anomalies and to derive a smooth curve by age. Beer's ordinary minimized fifth difference formula is such a technique and has been used in the construction of the previous U.S. decennial life tables (20) as well as annual life tables (21).

Population data are aggregated into 5-year age intervals except for those aged 100 and over, which are allocated into a single category. Values of P_x by single years of age were obtained by interpolation using Beer's formula. Beer's general formula adapted to calculate P_x is

$$P_{x+k} = C_{k,x-10} {}_5P_{x-10} + C_{k,x-5} {}_5P_{x-5} + C_{k,x} {}_5P_x + C_{k,x+5} {}_5P_{x+5} + C_{k,x+10} {}_5P_{x+10},$$

where P_{x+k} is the population aged $x+k$ ($k=0,1,2,3,4$), ${}_5P_x$ is the total population aged x to $x+k$, and C_{kx} is Beer's interpolation coefficient for the k^{th} fifth of the age interval x to $x+5$ applied to ${}_5P_x$. To interpolate single-year values from ${}_5P_0$, and ${}_5P_5$, the formula is slightly different. To obtain single-year values for these 5-year intervals, the formulas below are used.

$$P_{0+k} = C_{k,0} {}_5P_0 + C_{k,5} {}_5P_5 + C_{k,10} {}_5P_{10} + C_{k,15} {}_5P_{15} + C_{k,20} {}_5P_{20}$$

$$P_{5+k} = C_{k,0} {}_5P_0 + C_{k,5} {}_5P_5 + C_{k,10} {}_5P_{10} + C_{k,15} {}_5P_{15} + C_{k,20} {}_5P_{20}$$

Values for $C_{k,x}$ are shown in Table C.

Interpolating single-year values for D_x was conducted in a similar fashion as that for P_x . The Beer's coefficients $C_{k,x}$ for deaths are the same as for ${}_5P_x$ as shown in Table C. The difference is that when interpolating ages 5–9 and 10–14 years, a fictitious value for ${}_5D_0$ is used. Because of the mortality peak in infancy, the use of the observed ${}_5D_0$ does not yield values for D_x for ages 5–14 years that join smoothly with the numbers reported at ages 0–4 years. The fictitious value for ${}_5D_0$ is calculated such that

$$V = .4072 {}_5D_0^* + .2146 {}_5D_5 + .0080 {}_5D_{10} - .0896 {}_5D_{15} + .0328 {}_5D_{20},$$

where V is the sum of the deaths occurring at ages 2, 3, and 4 years and ${}_5D_0^*$ is the fictitious value for ${}_5D_0$. Solving for ${}_5D_0^*$ gives

$${}_5D_0^* = 2.45580 V - .59332 {}_5D_5 - .01965 {}_5D_{10} + .22004 {}_5D_{15} - .08055 {}_5D_{20},$$

${}_5D_0^*$ is then treated in the interpolation formula as if it were the actual number of deaths at ages 0–4 years. This modification produces a smooth transition from the observed values at ages under 5 years to the interpolated values at ages 5 years and over.

Table C. Interpolation coefficients based on Beer's ordinary formula for the subdivision of grouped data into fifths

Interpolated subgroup ($x+k$)	5-year age interval beginning with age				
	0	5	10	15	20
2 to 4 years					
2 years	+0.1924	+0.0064	+0.0184	-0.0256	+0.0084
3 years	+0.1329	+0.0844	+0.0054	-0.0356	+0.0129
4 years	+0.0819	+0.1508	-0.0158	-0.0284	+0.0115
5 to 9 years					
5 years	+0.0404	+0.2000	-0.0344	-0.0128	+0.0068
6 years	+0.0093	+0.2268	-0.0402	+0.0028	+0.0013
7 years	-0.0108	+0.2272	-0.0248	+0.0112	-0.0028
8 years	-0.0198	+0.1992	+0.0172	+0.0072	-0.0038
9 years	-0.0191	+0.1468	+0.0822	-0.0084	-0.0015
5x to 5x+5 years					
	5x-10	5x-5	5x	5x+5	5x+10
5x years	-0.0117	+0.0804	+0.1570	-0.0284	+0.0027
5x+1 year	-0.0020	+0.0160	+0.2200	-0.0400	+0.0060
5x+2 years	+0.0050	-0.0280	+0.2460	-0.0280	+0.0050
5x+3 years	+0.0060	-0.0400	+0.2200	+0.0160	-0.0020
5x+4 years	+0.0027	-0.0284	+0.1570	+0.0804	-0.0117
95 to 99 years					
	80	85	90	95	100
95 years	-0.0015	-0.0084	+0.0822	+0.1468	-0.0191
96 years	-0.0038	+0.0072	+0.0172	+0.1992	-0.0198
97 years	-0.0028	+0.0112	-0.0248	+0.2272	-0.0108
98 years	+0.0013	+0.0028	-0.0402	+0.2268	+0.0093
99 years	+0.0068	-0.0128	-0.0344	+0.2000	+0.0404

NOTE: An additional panel of interpolation coefficients is usually shown for the final age interval. This panel is not shown because the final age interval in the life table is open-ended and is not based on interpolated data.

SOURCE: See reference 20 in this report.

Central death rates at ages 2–4 years

The life table function q_x , denoting the ratio d_x/l_x , is the fraction or proportion of a group of persons at exact age x who are expected to die before attaining age $x + 1$. If m_x denotes the ratio d_x/L_x , commonly called the central death rate, then on the assumption of uniform distribution of deaths over the year at age x ,

$$q_x = \frac{2m_x}{2 + m_x}. \quad [1]$$

This approximation is appropriate when the life table is by single years of age. This formula is the basis of the calculation of probabilities of dying at ages 2–94. Completion of the calculations depends, therefore, on the ability to calculate the central death rate m_x at different ages. For this purpose, different methods are used at ages 2–4 years and at ages 5–94.

If D_x denotes the adjusted number of deaths in a population category at age x (in completed years) occurring in 2009–2011, and P_x denotes the population at age x in the middle of the period, then m_x is equivalent to

$$M_x = \frac{D_x}{3P_x}. \quad [2]$$

As previously noted, the populations used were those of April 1, 2010.

However, because the deaths occurring in a single year of age during 2009–2011 were drawn from three consecutive annual cohorts of the population, it was considered that the accuracy of these M_x values would be improved by replacing $3P_x$ in the denominator of formula [2] by the sum of the populations at ages $x - 1$, x , and $x + 1$. Thus, the formula becomes

$$M_x = \frac{D_x}{P_{x-1} + P_x + P_{x+1}}. \quad [3]$$

The combination of formulas [1] and [3] is equivalent to the single formula

$$q_x = \frac{D_x}{P_{x-1} + P_x + P_{x+1} + \frac{1}{2}D_x}, \quad [4]$$

which is used for $x = 2, 3$, and 4 , with values of D_x and P_x obtained by interpolation from data by five 5-year age intervals.

Probabilities of dying at ages 5–94

The combination of formulas [1] and [2] is equivalent to

$$q_x = \frac{D_x}{3P_x + \frac{1}{2}D_x}, \quad [5]$$

which is used for ages 5–94, with values of D_x and P_x obtained by interpolation from data by five 5-year age intervals.

Probabilities of dying at the oldest ages for the total, white, black, non-Hispanic white, and non-Hispanic black populations

Medicare data are used to supplement vital statistics data for the estimation of q_x at the oldest ages because these data are more accurate given that proof of age is required for enrollment in the Medicare program. Medicare data are used here to estimate the probability of dying for ages 66 and over for the total, white, black, non-Hispanic white, and non-Hispanic black populations.

The method described in this section consists of the following steps. First, vital statistics and Medicare death rates are blended in the age range 66–99. Second, a logistic model is used to smooth the blended death rates in the age range 85–99 and predict death rates for ages 100–120. Third, final resulting death rates, M_x , are converted to q_x .

For ages 66–94, vital statistics death rates, M_x^V , and Medicare death rates, M_x^M , are blended with a weighting process that gives gradually declining weight to vital statistics data and gradually increasing weight to Medicare data. For ages 95–99, M_x^M is used exclusively. Blended M_x is thus obtained as follows:

$$M_x = \frac{1}{30} [(95 - x)M_x^V + (x - 65)M_x^M], \quad [6]$$

when $x = 66, \dots, 94$, and $M_x = M_x^M$,
when $x = 95, \dots, 99$.

M_x^M is estimated as follows:

$$M_x^M = \frac{D_x^M}{P_x^M},$$

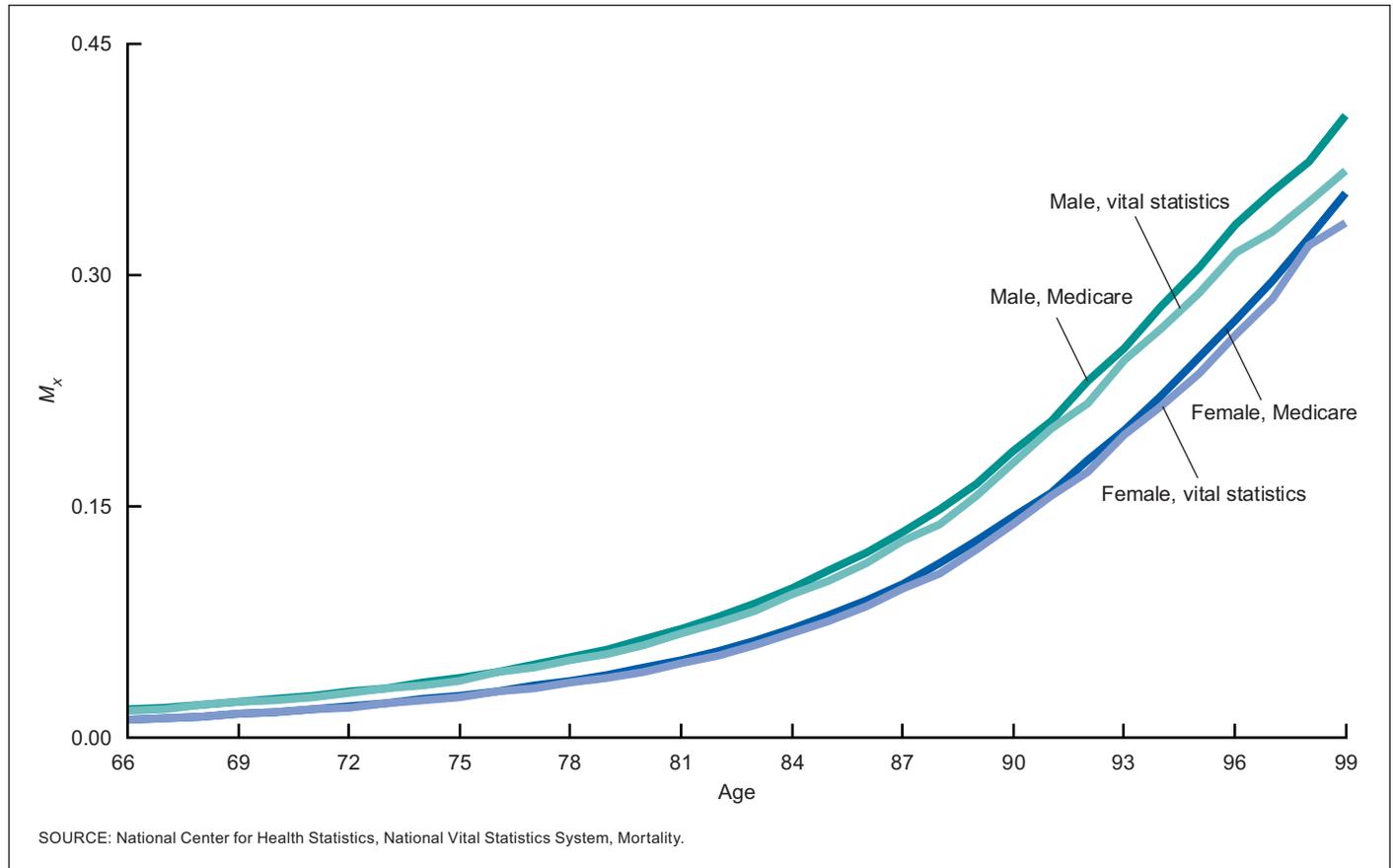
where D_x^M is the sum of age-specific Medicare death counts in years 2009, 2010, and 2011; and P_x^M is the sum of age-specific Medicare midyear population counts in years 2009, 2010, and 2011.

A logistic model proposed by Kannisto is then used to smooth M_x in the age range 85–99 and predict M_x in the age range 100–120 (3). The start of the modeled age range varies by race-, ethnicity-, and sex-specific population because it is a function of the age at which the rate of change in the age-specific death rates peaks. In current times, the rate of change in the age-specific death rate rises steadily up to approximately ages 80–85 or so and then begins to decline. As a result, it is difficult to model a large age span, such as 66–99, with one simple model without oversmoothing and thus altering the underlying mortality pattern observed in the population of interest (22). Further, the observed data for the age range 66–85 is reliable and robust, as indicated by the very close similarity between vital statistics and Medicare death rates, so it is unnecessary to model (smooth) the entire age span (66–99) (see Figure 1).

The Kannisto model is a simple form of a logistic model in which the logit of u_x (or the natural log of the odds of u_x) is a linear function of age, x (3). It is expressed as:

$$\ln \left[\frac{u_x}{1 - u_x} \right] = \ln(\alpha) + \beta x, \quad [7]$$

Figure 1. Comparison of age-specific death rates M_x from vital statistics and Medicare data in the total population, by sex: United States, 2009–2011



where u_x , the force of mortality (or the instantaneous death rate), is defined as:

$$u_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}}. \quad [8]$$

Because u_x is not directly observed but is closely approximated by m_x and $m_x = M_x$, then the logit of M_x is modeled instead. A maximum-likelihood generalized linear model estimation procedure is used to fit the following model in the age range 85–99:

$$\ln \left[\frac{M_x}{1 - M_x} \right] = \ln(\alpha) + \beta x. \quad [9]$$

Then, the estimated parameters are used to predict \bar{M}_x as:

$$\bar{M}_x = \frac{e^a e^{bx}}{1 + e^a e^{bx}}, \text{ or equivalently, } \bar{M}_x = \frac{e^{a+bx}}{1 + e^{a+bx}}, \quad [10]$$

where a and b are the predicted values of parameters $\ln(\alpha)$ and β , respectively, given by fitting model [9]. Estimated parameters and the starting age for the modeled age span by population for the period 2009–2011 are presented in [Table D](#).

Finally, the predicted probability of death, \bar{q}_x , for ages 85–120 is estimated by converting \bar{M}_x as:

$$\bar{q}_x = \frac{\bar{M}_x}{1 + \frac{1}{2} \bar{M}_x}. \quad [11]$$

Probabilities of dying at the oldest ages for the Hispanic population

As noted above, Medicare data are unreliable for the Hispanic population due to inconsistencies in the Medicare race and ethnicity classification system. As a result, it was necessary to use other methods to estimate mortality at the oldest ages for this population. Beyond age 80, mortality estimates based strictly on vital statistics for the Hispanic population are too low, despite correction for ethnic misclassification on the death certificate.

A consistent finding across diverse studies has been that Hispanic mortality in the adult and advanced ages varies between approximately 80% and 89% of that of the non-Hispanic white population (12,13,19,23). The Brass relational logit model takes advantage of the relationship between Hispanic and non-Hispanic white mortality previously identified and has been widely and successfully used to predict the mortality of one population relative to another at the older ages (24–27). Using the age-specific mortality pattern of the non-Hispanic white population as the standard, the Brass relational logit model is used to predict Hispanic mortality in the older ages. The standard is fit to Hispanic data in the age interval 45–80, and the predicted parameters are used to estimate the probabilities of death for ages 76–100. This method allows the relationship between the two populations in the younger ages to be carried over to the older ages (24–27).

Table D. Estimated parameters α and β used for predicting m_x and starting age of modeled age span: U.S. Decennial Life Tables, 2009–2011

Parameter	Total			White			Black			Non-Hispanic white			Non-Hispanic black		
	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female	Both sexes	Male	Female
Starting age															
$\ln(\alpha)$	0.1271821	0.1266491	0.1337931	0.1306361	0.1317536	0.1369603	0.0994343	0.0929695	0.1074015	0.1303803	0.1314313	0.136874	0.098664	0.0915774	0.106704
(SE)	(0.0007173)	(0.0019147)	(0.0005557)	(0.0009834)	(0.0024514)	(0.0005429)	(0.0015869)	(0.0041612)	(0.0013916)	(0.0009643)	(0.0024298)	(0.000598)	(0.0015445)	(0.0044922)	(0.0013633)
β	-13.65263	-13.39578	-14.37753	-13.96334	-13.85908	-14.66253	-11.12418	-10.29667	-11.96706	-13.93761	-13.82662	-14.65361	-11.04729	-10.15807	-11.89717
(SE)	(0.0679395)	(0.1809904)	(0.0531749)	(0.0942997)	(0.233792)	(0.0519785)	(0.146101)	(0.3827591)	(0.1279522)	(0.0925101)	(0.2317666)	(0.0574443)	(0.1421353)	(0.4155488)	(0.1253189)

NOTE: SE is standard error.

SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

Table E. Estimated Brass relational logit model parameters α and β for the Hispanic-origin population: U.S. Decennial Life Tables, 2009–2011

Parameter	Total (SE)	Male (SE)	Female (SE)
α	0.9922375 (0.0047209)	0.9868890 (0.0074861)	1.0176780 (0.0073409)
β	-0.2897252 (0.0164864)	-0.2609211 (0.0278626)	-0.2267147 (0.0353827)

NOTE: SE is standard error.

SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

The Brass relational logit model expresses the age-specific mortality pattern of a population of interest as a function of the age-specific mortality pattern of a standard population and is expressed as:

$$\bar{Y}_x = \alpha + \beta Y_x^S, \quad [12]$$

where \bar{Y}_x is the predicted logit of the probability of death, q_x in the population of interest, that is,

$$\text{logit}[q_x] = \ln \left[\frac{q_x}{1 - q_x} \right].$$

Y_x^S is the logit of the probability of death in the standard population, q_x^S , that is,

$$\text{logit}[q_x^S] = \ln \left[\frac{q_x^S}{1 - q_x^S} \right].$$

α is the predicted parameter that measures the level of mortality of the population of interest relative to the standard population, and β is the predicted parameter that measures the slope of the mortality function of the population of interest relative to the standard population (4,25–27). Table E shows values of predicted α and β and their standard errors.

A maximum-likelihood generalized linear model estimation procedure is used to fit equation [12] in the age range 45–80. The resulting predicted parameters α and β are then used to estimate the predicted probability of death for ages 76–120 in the Hispanic population.

Predicted \bar{q}_x is estimated by transforming its logit, \bar{Y}_x , back as:

$$\bar{q}_x = \frac{\exp[\bar{Y}_x]}{1 + \exp[\bar{Y}_x]} = \frac{\exp[\alpha + \beta Y_x^S]}{1 + \exp[\alpha + \beta Y_x^S]}. \quad [13]$$

To ensure a smooth transition from vital q_x^V and predicted, \bar{q}_x , the two were blended from ages 76 to 80 with a graduating process as:

$$q_x = \frac{1}{6} [(81 - x)q_x^V + (x - 75)\bar{q}_x], \quad [14]$$

when $x = 76, \dots, 80$.

Calculation of remaining life table functions for all groups

Survivor function (l_x)

The life table radix, l_0 , is set at 100,000. The number of survivors remaining at exact age x is calculated as:

$$l_x = l_{x-1} (1 - q_{x-1}). \quad [15]$$

Decrement function (d_x)

The number of deaths occurring between ages x and $x + 1$ is calculated from the survivor function

$$d_x = l_x - l_{x+1} = l_x q_x. \quad [16]$$

Person-years lived (L_x)

Person-years lived is calculated assuming that the survivor function declines linearly between ages x and $x + 1$. This gives the formula

$$L_x = \frac{1}{2} (l_x + l_{x+1}) = l_x - \frac{1}{2} d_x. \quad [17]$$

For ages under 1 year, L_x should be expressed as a quantity in a fractional year, that is, L_x should be calculated by multiplying the equation by 1/365 for age 0–1 day, by 6/365 for 1–7 days, by 21/365 for 7–28 days, and by 337/365 for 28–365 days.

Person-years lived at and above age x (T_x)

T_x is calculated by summing L_x values at and above age x :

$$T_x = L_x + T_{x+1}, \text{ with } T_{\text{end}} = L_{\text{end}}. \quad [18]$$

Life expectancy at age x , (e_x)

Life expectancy at exact age x is calculated as:

$$e_x = \frac{T_x}{l_x}. \quad [19]$$

Finally, all life table functions are estimated through ages 120 years but truncated at age 109 years for publication.

Variances and standard errors of the probability of dying and life expectancy

The mortality data on which the life tables are based is not affected by sampling error because it is based on complete counts of deaths and, as a result, variances and standard errors reflect only stochastic (random) variation. While measurement error, such as age and race and ethnicity misclassification, is known to affect mortality estimates, it is not considered in the calculation of the variances or standard errors of the life table functions (11). Because the life tables for the United States are based on relatively large numbers of deaths, the variances and standard errors presented are rather small.

The methods used to estimate the variances of q_x and e_x are based on Chiang (28) with some necessary modifications due to the use of statistical modeling for smoothing and prediction of old-age death rates. Based on the assumption that deaths are binomially distributed, Chiang proposed the following equation for the variance of q_x :

$$\text{Var}(q_x) = \frac{q_x^2(1 - q_x)}{D_x}, \quad [20]$$

where D_x is the age-specific number of deaths. For the total, white, black, non-Hispanic white, and non-Hispanic black populations, this equation is used to estimate $\text{Var}(q_x)$ throughout the age span with a modification where, for ages under 66, D_x is the deaths from vital statistics data, smoothed by interpolation and adjusted for the number of deaths with age not stated and

for the non-Hispanic white and non-Hispanic black populations adjusted for race and Hispanic-origin misclassification. For ages 66 and over, D_x was obtained by treating the population as a cohort population and calculated from q_x because blended vital statistics and Medicare data were used for estimation (29):

$$P_x = \frac{\left(\frac{P_{x-1} - 0.5D_{x-1}}{3}\right)(2 - q_x)}{2},$$

$$D_x = \frac{3q_x P_x}{1 - 0.5q_x}.$$

Note that D_x is the number of deaths in a 3-year data collection period (2009–2011), and P_x is the population at age x in the middle of this period.

For the Hispanic population, equation 20 is used for ages 0–75. Because q_x for ages 81 and over is predicted based on the Brass relational logit model, the Delta method is used to approximate its variance as follows:

$$\text{Var}(\bar{q}_x) = \left(\frac{\exp(\alpha + \beta Y_x^S)}{(\exp(\alpha + \beta Y_x^S) + 1)^2}\right)^2 \cdot \text{Var}(\alpha + \beta Y_x^S). \quad [21]$$

For ages 76–80, the variance of q_x is calculated as:

$$\frac{1}{6} [(81 - x) \cdot \text{Var}(q_x) + (x - 75) \cdot \text{Var}(\bar{q}_x)]. \quad [22]$$

Standard error of q_x :

$$SE(q_x) = \sqrt{\text{Var}(q_x)}. \quad [23]$$

For all groups, variances of the life expectancies for ages 0–109 are estimated using Chiang's equation:

$$\text{Var}(e_x) = \frac{\sum_{x=0}^{x=109} l_x^2 \cdot [(1 - 0.5) + e_{x+n}]^2 \cdot \text{Var}(q_x)}{l_x^2}. \quad [24]$$

Standard error of e_x :

$$SE(e_x) = \sqrt{\text{Var}(e_x)}. \quad [25]$$

Conclusions

The methodology presented in this report details the estimation of the 2009–2011 decennial life tables and the standard errors of the probability of dying and life expectancy. Life tables were produced for the total, male, female, white, white male, white female, black, black male, black female, Hispanic, Hispanic male, Hispanic female, non-Hispanic white, non-Hispanic white male, non-Hispanic white female, non-Hispanic black, non-Hispanic black male, and non-Hispanic black female populations (1).

References

1. Arias E, Miniño A, Curtin SC, Tejada-Vera B. U.S. decennial life tables for 2009–2011, United States life tables. National Vital Statistics Reports; vol 69 no 8. Hyattsville, MD: National Center for Health Statistics. 2020. [Forthcoming].
2. Arias E. United States life tables, 2008. National Vital Statistics Reports; vol 61 no 3. Hyattsville, MD: National Center for Health Statistics. 2012. Available from: https://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_03.pdf.
3. Thatcher AR, Kannisto V, Vaupel JW. The force of mortality at ages 80 to 120. Odense, Denmark: Odense University Press. 1998.
4. Arias E. United States life tables by Hispanic origin. National Center for Health Statistics. Vital Health Stat 2(152). 2010. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_152.pdf.
5. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. Fed Regist 62(210):58782–90. 1997. Available from: <https://www.resolutionmineeis.us/sites/default/files/references/omb-standards-revisions-1997.pdf>.
6. Ingram DD, Parker JD, Schenker N, Weed JA, Hamilton B, Arias E, Madans JH. United States Census 2000 population with bridged race categories. National Center for Health Statistics. Vital Health Stat 2(135). 2003. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_135.pdf.
7. Office of Management and Budget. Race and ethnic standards for federal statistics and administrative reporting. Statistical Policy Directive 15. 1977. Available from: <https://wonder.cdc.gov/WONDER/help/populations/bridged-race/Directive15.html>.
8. Greville TNE. United States life tables: 1969–1971, U.S. decennial life tables for 1969–1971; vol 1, no 1. DHEW Pub. No. HRA 75–1150. Rockville, MD: U.S. Department of Health, Education, and Welfare. Health Resources Administration. National Center for Health Statistics. 1975.
9. Bell FC, Miller ML. Life tables for the United States Social Security Area 1900–2100. Baltimore, MD: Social Security Administration, Office of the Chief Actuary. SSA Pub. No. 11–11536. 2005.
10. Research Data Assistance Center. Introduction to the use of Medicare data for research. Minneapolis, MN: University of Minnesota School of Public Health. 2004.
11. Arias E, Heron M, Hakes JK. The validity of race and Hispanic-origin reporting on death certificates in the United States: An update. Vital Health Stat 2(172). Hyattsville, MD: National Center for Health Statistics. 2016. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_172.pdf.
12. Arias E, Schauman WS, Eschbach K, Sorlie PD, Backlund E. The validity of race and Hispanic origin reporting on death certificates in the United States. National Center for Health Statistics. Vital Health Stat 2(148). 2008. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_148.pdf.
13. Arias E, Eschbach K, Schauman WS, Backlund EL, Sorlie PD. The Hispanic mortality advantage and ethnic misclassification on U.S. death certificates. Am J Public Health 100(Suppl 1):S171–7. 2010.
14. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2007 period linked birth/infant death data set. National Vital Statistics Reports; vol 59 no 6. Hyattsville, MD: National Center for Health Statistics. 2011.
15. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2008 period linked birth/infant death data set. National Vital Statistics Reports; vol 60 no 5. Hyattsville, MD: National Center for Health Statistics. 2012.

16. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2009 period linked birth/infant death data set. *National Vital Statistics Reports*; vol 61 no 8. Hyattsville, MD: National Center for Health Statistics. 2013.
17. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2010 period linked birth/infant death data set. *National Vital Statistics Reports*; vol 62 no 8. Hyattsville, MD: National Center for Health Statistics. 2013.
18. National Center for Health Statistics. User guide to the 2011 period linked birth/infant death public use file. *National Vital Statistics System*.
19. Turra CM, Elo IT. The impact of salmon bias on the Hispanic mortality advantage: New evidence from Social Security data. *Popul Res Policy Rev* 27(5):515–30. 2008.
20. Shryock HS, Siegel JS, Larmon EA. The methods and materials of demography, vol 2. U.S. Census Bureau. Washington, DC: U.S. Government Printing Office. 1971.
21. Arias E, Xu J. United States life tables, 2017. *National Vital Statistics Reports*; vol 68 no 7. Hyattsville, MD: National Center for Health Statistics. 2019.
22. Andreev KF, Bourbeau RR. Frailty modeling of Canadian and Swedish mortality at adult and advanced ages. Silver Spring, MD: Population Association of America. 2007.
23. Elo IT, Turra CM, Kestenbaum B, Fergusson BR. Mortality among elderly Hispanics in the United States: Past evidence and new results. *Demography* 41(1):109–28. 2004.
24. Preston SM, Heuveline P, Guillot M. *Demography: Measuring and modeling population processes*. Oxford: Blackwell Publishers. 2001.
25. Brass W. On the scale of mortality. In: Brass W, editor. *Biological aspects of demography*. London: Taylor and Francis. 99–110. 1971.
26. Himes CL, Preston SH, Condran GA. A relational model of mortality at older ages in low mortality countries. *Popul Stud* 48(2):269–91. 1994.
27. Preston SH, Elo IT. Black mortality at very old ages in official U.S. life tables: A skeptical appraisal. *Popul Dev Rev* 32(3):557–65. 2006.
28. Chiang CL. *The life table and its applications*. Malabar, FL: Krieger Publishing. 1984.
29. Wei R, Curtin LR, Arias E, Anderson RN. U.S. decennial life tables for 1999–2001: Methodology of the United States life tables. *National Vital Statistics Reports*; vol 57 no 4. Hyattsville, MD: National Center for Health Statistics. 2008.

**U.S. DEPARTMENT OF
HEALTH & HUMAN SERVICES**

Centers for Disease Control and Prevention
National Center for Health Statistics
3311 Toledo Road, Room 4551, MS P08
Hyattsville, MD 20782-2064

FIRST CLASS MAIL
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

For more NCHS NVSRs, visit:
<https://www.cdc.gov/nchs/products/nvsr.htm>.



National Vital Statistics Reports, Vol. 69, No. 10, August 31, 2020

Contents

Abstract	1
Introduction	1
Methods	2
Data used for calculating life table values	2
Preliminary adjustment of the data	3
Calculation of the probability of dying, ${}_nq_x$	4
Calculation of remaining life table functions for all groups	9
Variances and standard errors of the probability of dying and life expectancy	9
Conclusions	10
References	10

Acknowledgments

The authors are grateful for the reviews and comments provided by Robert N. Anderson, Mortality Statistics Branch (MSB), Division of Vital Statistics (DVS); Isabelle Horon, DVS; and Amy Branum, Office of the Director. The authors thank Kenneth Kochanek, MSB, for content review. The report was edited and produced by NCHS Office of Information Services, Information Design and Publishing Staff: Nora Castro edited the report; typesetting was done and graphics were produced by Simon McCann.

Suggested citation

Arias E, Curtin SC, Tejada-Vera B. U.S. decennial life tables for 2009–2011: Methodology of the United States life tables. National Vital Statistics Reports; vol 69 no 10. Hyattsville, MD: National Center for Health Statistics. 2020.

Copyright information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

National Center for Health Statistics

Brian C. Moyer, Ph.D., *Director*
Amy M. Branum, Ph.D., *Acting Associate
Director for Science*

Division of Vital Statistics

Steven Schwartz, Ph.D., *Director*
Isabelle Horon, Dr.P.H., *Acting Associate
Director for Science*